Title
The effects of exposure to differing amounts of misinformation and source credibility perception on source monitoring and memory accuracy.

Journal
Psychology of Consciousness: Theory, Research, and Practice, 4(4)

ISSN
2326-5523

Authors
Pena, Michelle M
Klemfuss, J Zoe
Loftus, Elizabeth F
et al.

Publication Date
2022-12-21

DOI
10.1037/cns0000137

Peer reviewed

Psychology of Consciousness: Theory, Research, and Practice

The Effects of Exposure to Differing Amounts of Misinformation and Source Credibility Perception on Source Monitoring and Memory Accuracy

Michelle M. Pena, J. Zoe Klemfuss, Elizabeth F. Loftus, and Amelia Mindthoff

Online First Publication, October 23, 2017. http://dx.doi.org/10.1037/cns0000137

Psychology of Consciousness: Theory, Research, and Practice

2326-5523/17/$12.00 http://dx.doi.org/10.1037/cns0000137

The Effects of Exposure to Differing Amounts of Misinformation and Source Credibility Perception on Source Monitoring and Memory Accuracy

Michelle M. Pena
Florida International University

J. Zoe Klemfuss
Florida International University and University of California, Irvine

Elizabeth F. Loftus
University of California, Irvine

Amelia Mindthoff
Florida International University

Although it is well known that exposure to misinformation after an event can alter memory, less known are the effects of being presented with different *amounts* of misinformation. The present study examined (a) how exposure to different amounts of misinformation affects memory, (b) how sensitively individuals monitor the accuracy of a (mis)information source, (c) whether perceived credibility of the misinformation source mediates the relations between misinformation exposure and memory accuracy, (d) whether perceived source credibility is associated with improved source monitoring, and (e) how exposure to different amounts of misinformation affects the ability to accurately assess one's own memory performance. Participants watched a mock crime video, were exposed to a misleading narrative about the video containing 20%, 50%, or 80% misinformation, completed a memory test, and rated the credibility of the misinformation source and their own memory performance. Receiving more misinformation decreased memory accuracy. Interestingly, receiving more misinformation also led subjects to become more skeptical of the credibility of the narrative, dampening the negative effect of misinformation on memory accuracy. In addition, individuals' perceptions of the source's credibility and source monitoring accuracy were negatively associated. Lastly, participants' performance estimates and confidence were well calibrated to their actual performance, except when they were misled, supporting the idea that misinformed responses are more difficult to monitor. Participants also tended to overestimate their accuracy, particularly when they performed poorly.

Keywords: misinformation, source credibility, memory accuracy, metacognition

Memory is a constant restructuring process, one that consists of incorporating new information into old memories and interpreting new information in light of what is already known (see Clifasefi, Garry, & Loftus, 2007; Schacter, 2001, for reviews). Researchers have reported that this restructuring process can result in memory errors when the new information presented is incorrect (Belli, 1989; Dodd & Bradshaw, 1980; Loftus, 1975; Loftus, 1979). These findings have led to further explorations of the processes that may be at play within the misinformation effect in efforts to provide possible explanations as to *why* individuals incorporate postevent misinformation into memory. Understanding the cognitive and metacognitive aspects of memory, such as the ability to monitor postevent information for accuracy, to accurately distinguish an experienced memory from

Michelle M. Pena, Department of Psychology, Florida International University; J. Zoe Klemfuss, Department of Psychology, Florida International University, and Department of Psychology and Social Behavior, University of California, Irvine; Elizabeth F. Loftus, Department of Psychology and Social Behavior, University of California, Irvine; Amelia Mindthoff, Department of Psychology, Florida International University.

Correspondence concerning this article should be addressed to J. Zoe Klemfuss, Department of Psychology and Social Behavior, University of California, Irvine, 4201 Social and Behavioral Sciences Gateway, Irvine, CA 92697. E-mail: jklemfus@uci.edu

postevent misinformation, and to correctly assess one's own memory performance, has important implications for our theoretical understanding of how the misinformation effect occurs.

In the present study, we examined contextual and individual difference effects on resistance to false information provided after exposure to a mock crime. Specifically, we explored (a) whether increasing the proportion of misinformation to neutral information negatively impacts memory accuracy, (b) how sensitively individuals monitor the accuracy of a (mis)information source, (c) whether perceived credibility of the misinformation source mediates the relations between misinformation exposure and memory accuracy, (d) whether perceived source credibility leads to improved source monitoring, and (e) the accuracy of participants' self- evaluations of memory performance. By collectively examining these issues, the present study was designed to provide additional support for the negative impact of misinformation on memory accuracy, but more importantly, the design allowed for the investigation of factors (perceived source credibility, source monitoring accuracy) that may be heightened in the face of high rates of misinformation and aid in the resistance to that misinformation. Further, we were able to determine how effectively participants are able to monitor their memory performance when misinformed.

Misinformation and Memory Performance

Provision of false information between a to-be-remembered event and recall of that event is a common source of memory error (Loftus & Hoffman, 1989). The misinformation effect has the potential to influence memory of an original event, which can lead to decreased accuracy at recall (Belli, 1989; Loftus & Hoffman, 1989; Loftus, Miller, & Burns, 1978; Weingardt, Loftus, & Lindsay, 1995). One cognitive explanation for the misinformation effect is suggested by the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), which claims that every memory, including memory for misinformation, has a source, which is evaluated by the individual at the time of retrieval. Sometimes misinformation is recalled instead of what was actually witnessed simply because the source of the misinformation is confused as being from the original event, an error referred to as a source misattribution error (Belli, Lindsay, Gales, & McCarthy, 1994; Zaragoza & Lane, 1994). In line with an integrated model of the misinformation effect, research suggests that source monitoring is a cognitive process that can be affected by social factors such as the perceived credibility of the postevent source (Skagerberg & Wright, 2009).

Source Credibility and Memory Performance

People who are given warnings about the credibility of the source of postevent information (e.g., that the source is not credible, or that the source is someone who may have motivation to deceive) scrutinize the features of the source more than those who are not given a warning. This additional scrutiny allows the warned group to more effectively monitor their memories and to dismiss misleading items (e.g., Echterhoff, Hirst, & Hussy, 2005; Frost, Ingraham, & Wilson, 2002; Skagerberg & Wright, 2009). However, although research suggests that individuals are capable of making credibility judgments from implicit warnings (e.g., the source is the driver who caused the accident being described), it is less clear whether individuals are able to make such judgments based solely on the content of postevent misinformation without alluding to any specific characteristic of the source that may affect credibility judgments. It is important to examine cues to credibility that are intrinsic to the actual quality of the information received, as they are in many cases in the real world.

Discrepancy Detection and Memory Performance

Existing literature on the principle of discrepancy detection, proposed by Tousignant, Hall, and Loftus (1986), suggests that individuals may be capable of accurately assessing the (in-)accuracy of postevent narratives containing misinformation without any cues other than the content of the narrative. When individuals are able to detect discrepancies between an event and postevent misinformation, they are better able to reject misinformation (Tousignant et al., 1986). For example, increasing suggestive details may actually inhibit children from devel-

oping images to false memories because they are more likely to detect the discrepancies (Otgaar, Candel, Scoboria, & Merckelbach, 2010). In addition, Loftus (1979) found that providing individuals with blatant misinformation made them more resistant to misinformation. In the present study, we examined whether increasing the proportion of misinformation to neutral information in a postevent narrative would increase discrepancy detection by offering more opportunities to note discrepancies. To our knowledge, the present study is the first to investigate whether individuals can correctly assess the (in)accuracy of postevent narratives varying in the degree of misinformation without any previous experience with the information source or any blatant credibility cues.

Metacognition and Memory Performance

Metacognition may also play an important role in explaining the misinformation effect. Specifically, the metacognitive concept of feeling of knowing (FOK) refers to a self-prediction regarding performance on a knowledge-based task, which may affect what an individual reports when asked to recall an event. Previous research investigating the relation between an individual's FOK and actual performance on an exam demonstrates a tendency to overestimate one's performance and abilities (Falchikov & Boud, 1989). For example, Fischhoff, Slovic, and Lichtenstein (1977) found that when participants were given general-knowledge questions and subsequently asked to rate how confident they were in their responses, they were consistently overconfident in their responses. However, the majority of this research has focused on calibrating accuracy and confidence in terms of knowledge-based tasks in learning contexts (See Schwartz, 1994 for review). Less is known about whether individuals exhibit the same tendency to overestimate on tasks geared more toward memory accuracy.

Additionally, some research has examined the differences between misinformed and control subjects on their accuracy-confidence calibration (Tomes & Katz, 2000; Weingardt, Leonesio, & Loftus, 1994). For example, Weingardt and colleagues found that individuals who were misled endorsed greater confidence in their incorrect responses compared with those individuals who were not misled: Individuals who received misinformation experienced a high FOK when presented with a misinformation response option and therefore reported a high confidence judgment when choosing the misinformation response compared with individuals who chose the misled response but never received it. The present study aimed to address the relations between confidence judgments and accuracy performance by examining how confidence ratings differed among different types of responses (i.e., correct, incorrect, and misinformed responses) and how global performance estimates varied by actual performance.

Overview of the Present Study

We addressed the relations among misinformation, perceived source credibility, and memory accuracy. Participants watched a video of a mock crime and then received a misleading written narrative containing low (20%), medium (50%), or high (80%) proportions of misinformation. Participants then completed a memory test related to the video and answered questions regarding the narrative's credibility and their performance on the memory questions. This between-subjects design allowed us to test direct relations between amount of misinformation and memory accuracy, as well as the mediational effect of perceived credibility and relations with source monitoring accuracy. We were also able to examine whether individuals are capable of accurately estimating their performance and the effect of receiving different quantities of misinformation on the calibration of performance estimates and confidence with respect to actual performance.

We hypothesized that individuals exposed to larger proportions of misinformation would answer a smaller proportion of questions regarding misinformation items correctly and a larger proportion of those questions in line with the misinformation compared with individuals exposed to smaller proportions of misinformation. Our second hypothesis was that as misleading information increased in proportion to nonmisleading information, participants would perceive the source as less credible. Relatedly, we hypothesized that this change in perceived source credibility would mediate the relations between misinformation exposure and memory accuracy. Specifically, based on discrepancy detection literature, we hypothesized that a de-

crease in perceived credibility would result in a dampening of the misinformation effect. Fourth, based on the source monitoring view of misinformation resistance (Chambers & Zaragoza, 2001; Echterhoff et al., 2005), we hypothesized that perceived source credibility and source monitoring accuracy would be associated. Finally, we hypothesized that participants would overestimate their performance on the memory questions, especially if they were highly inaccurate, based on past research showing individuals' tendency to overestimate their performance, particularly if they are low performing (Falchikov & Boud, 1989).

Method

Participants

The participants were 142 students at a large southeastern university who were recruited via the university's psychology research participation system in exchange for course credit. Of these participants, six were excluded because of a recording equipment malfunction, six because they answered the attention-check question incorrectly, and three because they decided they did not want their data to be used. The final sample of 127 participants was 69% female ($N = 88$) and representative of the community from which they were drawn (65% Hispanic, 15% African American, 11% Caucasian, 4% Asian American, and 5% other). They ranged in age from 18 to 39 years ($M = 21.21$, $SD = 3.28$).

Materials and Procedure

Participants completed all study procedures via an online survey on individual computers in a computer lab on campus. As many as 12 participants were tested per session. Online consent and all study procedures were administered via the Qualtrics survey program and study participation lasted 60 to 80 min.

Mock crime video. After providing online consent, participants viewed a brief (six and a half minute) silent video of a mock crime, which has been successfully used in previous related research (e.g., Takarangi, Parker, & Garry, 2006). The video depicts an electrician who enters a house to fix various items such as the oven and a light fixture, but while doing so he also rummages through the owner's personal belongings and steals some of their possessions, such as a pair of earrings and a compact disk. After viewing the video, participants were taken to a new survey page and were given a 12-min filler task in which they were asked to answer trivia questions. Once the 12 min expired, they were automatically advanced to the next section of the study where they received further instructions.

Misleading narrative. Participants were informed that a previous set of participants had written brief narratives about the robbery depicted in the video and that one of those narratives would be presented to them to refresh their memory. They were shown a photograph and basic demographic information about the alleged participant whose narrative they were reading. This "previous participant" was actually an artificial profile composed of a randomly selected head shot chosen from a sample of eight, which was retrieved via a publically available online database of faces. Participants were also presented with artificially created demographic information for the eight profiles designed to loosely match the demographics of the population from which the sample was drawn: 21 year-old Caucasian male, 20 year-old Caucasian male, 21 year-old Hispanic male, 20 year-old Hispanic male, 21 year-old Caucasian female, 20 year-old Caucasian female, 21 year-old Hispanic female, and 20 year-old Hispanic female. The goal was to create believable profiles of information sources (alleged previous participants) that participants could later evaluate for credibility. To gauge whether participants believed that a previous participant had actually written the narrative, at the conclusion of the study they were asked, "Did you suspect that the profile and narrative you read were created by the research team instead of another participant? If so, what made you suspect?"

Participants were randomly assigned to receive a narrative with a low (20%), medium (50%), or high (80%) proportion of misleading information. Each narrative was composed of a total of 40 sentences and so those in the 20% condition received a narrative with eight misleading sentences, the 50% condition received a narrative with 20 misleading sentences, and the 80% condition received a narrative with 32 misleading sentences. The misleading sentences were randomized so that the specific sentences

that contained misleading information varied across participants even within the same condition. The pieces of misleading information were changes made to details about objects seen in the video; information was not added. The nonmisleading sentences were neutral to avoid strengthening the participant's memory trace for the target items. For example, a misleading item was "He noticed the dresser was covered with a *brown* cloth" and the neutral version of that item was "He noticed the dresser was covered with a cloth." Each of the 40 sentences was displayed to the participants one at a time. There was no time restriction per line so that the participants could take however much time they needed to read each sentence. After reading the entire narrative, participants completed a 5-min filler math test. After this, participants were automatically advanced to a new page where they answered questions about the mock crime video.

Memory questions. All participants were asked the same set of 40 three-option forced-choice questions about the mock crime video (e.g., Eric noticed that the dresser was covered with a cloth) in random order. For each memory question, participants could choose the correct answer, meaning, the option corresponding to what was actually portrayed in the video (e.g., navy), an incorrect answer that for some participants reflected a piece of misinformation provided in the narrative (e.g., brown), or an incorrect answer, corresponding to information that was not portrayed to any of the participants in the video or the narrative (e.g., yellow). The response choices were presented simultaneously with the order randomized. For each item participants were subsequently asked how confident they were of their answer, from one, meaning not very confident, to five, meaning very confident. Finally, after each item they were also asked the source of their answer (i.e., the video, the narrative, both, or neither). Participants were also asked an attention-check question to ensure that they actually read each sentence from the narrative presented to them. The attention-check question asked, "What is the answer?" to which they were instructed while reading the narrative to answer, "potato chips."

Because participants were exposed to differing amounts of misinformation, they also differed in the number of memory questions to which they could provide a misinformed response. We refer to these questions as "misinformation questions." To account for varying exposure to misinformation, we calculated the proportions of correct, incorrect, and mislead responses to misinformation questions for each participant. The denominator for each proportion score was the total number of misinformation questions asked. We used these proportion scores as our memory variables in all relevant analyses. For example, a participant in the 20% misinformation condition who answered six misinformation questions correctly would receive a proportion correct score of 6/8 = .75, whereas a participant in the 80% misinformation condition who similarly answered 6 misinformation questions correctly would receive a proportion correct score of 6/32 = .19.

Questionnaire. Following the questions about the mock crime video, participants were given a 28-item questionnaire. The questionnaire examined the participant's metacognition and perceived source credibility. Participants were first asked questions concerning their cognitive monitoring, which is the online monitoring of one's own cognitive processes. Specifically, participants were asked four questions about the strength of their memories (e.g., "How good is your memory for the video you just watched?"). All questions were answered on a 5-point Likert scale with lower numbers indicating weaker memory strength and higher numbers indicating stronger memory strength. These four items reached an acceptable level of consistency (Cronbach's $\alpha = .73$) and thus were averaged to create a memory strength variable. Participants were asked an additional three cognitive monitoring questions specifically aimed to gauge how well they thought they did on the memory questionnaire (e.g., "How many questions do you think you got *wrong*?"). Participants' answers to these cognitive monitoring questions were then directly compared to their actual performance.

Participants were then questioned about the credibility of the (mis)information source (e.g., "How good was the *quality* of the narrative?" "How *familiar* is the other participant with the video?"). Specifically, participants answered eight questions concerning the perceived credibility of the narrative and narrator on a 5-point Likert scale with lower numbers indicating lower credibility and higher numbers indicating

higher credibility. The eight items were consistent (Cronbach's $\alpha = .83$) and were averaged to create a perceived credibility variable. These items included helpfulness, quality, (in)accuracy, narrator familiarity with the video, believability, honesty, and expected memory performance if presented with a different narrative. They were also asked to indicate what percent of the narrative was accurate from 0% to 100%. Lastly, participants were asked basic demographic questions. After participants finished the questionnaire, they were debriefed and were given the opportunity to rewatch the video to refresh their memory.

Results

Preliminary Analyses

Participants were roughly evenly distributed across study conditions ($N = 39$ in the high misinformation condition, $N = 44$ in the medium misinformation condition, $N = 44$ in the low misinformation condition). There were no significant differences in age, gender, ethnicity, or year in college between the misinformation conditions. The eight false participant profiles were approximately evenly distributed across conditions such that a chi-square analysis revealed no significant differences in profile presentation across conditions. Approximately one third of participants indicated that they suspected the profiles to be false (35%). Interestingly, of these, half suspected the profile solely because they noted inaccuracies in the narrative allegedly provided by the person described in the false profile (17% of the total sample). Participants were retained in the sample regardless of their responses to the skepticism question given that these responses did not vary by condition, $\chi^2(6, N = 127) = 7.06$, $p = .315$, and were unrelated to the proportion of memory questions they answered correctly, $F(3, 123) = .23$, $p = .873$.

Next, we present analyses addressing our key research questions. First, we examine the effects of misinformation condition on proportions of correct, incorrect, and misled responses to misinformation questions. Second, we address our primary hypothesis regarding the mediating role of perceived source credibility in the relations between misinformation and memory performance via Andrew Hayes' PROCESS tool for SPSS (http://www.afhayes.com). Third, we probe the links between perceived source credibility and source monitoring accuracy. Finally, we present analyses linking perceived and actual memory performance.

Correct, Incorrect, and Misinformed Responses by Misinformation Group

Table 1 presents means and standard deviations for the proportions of correct, incorrect, and misled responses participants provided to misleading questions, by misinformation condition. We found (a) significant differences between conditions in the proportion of correct, $F(2, 124) = 5.35$, $p = .006$, $\eta_p^2 = .079$, and misled responses, $F(2, 124) = 7.77$, $p = .001$, $\eta_p^2 = .111$, provided, and (b) no significant condition effects on the proportion of incorrect responses provided to misleading questions, $F(2, 124) = 0.02$, $p = .984$, $\eta_p^2 = .000$. Tukey's post hoc comparisons revealed that participants in the 80% misinformation condition provided a smaller proportion of correct responses and a higher proportion of misled responses to misinformation items compared with participants in the 20% misinformation condition, $ps < .01$. We found no significant differences between the 50% condition and either of the other conditions, $ps > .05$.

Source Credibility as a Mediator Between Misinformation Group and Memory Accuracy

Our primary research question involved the mediating role of perceived source credibility between receipt of misinformation and memory accuracy, which we have defined as the proportion of misleading questions answered correctly. In the first set of models, the independent

Table 1
Means and Standard Deviations for Proportions of Correct, Incorrect, and Misled Responses Provided to Misleading Items, by Misinformation Condition

	Misinformation condition		
Response type	80%	50%	20%
Correct	.63 (.13)*	.66 (.16)	.74 (.18)*
Misled	.30 (.13)**	.26 (.14)	.19 (.13)**
Incorrect	.07 (.05)	.07 (.08)	.07 (.09)

* $p = .006$. ** $p = .001$.

variable was misinformation condition dummy coded to compare high versus low, in the second set of models, misinformation condition was dummy coded to compare medium versus low. We found a significant indirect effect of the high versus low misinformation condition on the proportion of correct responses through perceived credibility, $b = .04$, $p = .021$, bias corrected and accelerated (BCa) 95% confidence interval (CI) [.01, .09], and through estimated percentage of the narrative that was correct, $b = .06$, $p = .006$, BCa 95% CI [.03, .11]. A similar pattern of effects was evident when we compared the medium and low misinformation conditions, except that only the indirect effect through estimated percentage of the narrative that was correct was statistically significant, $b = .06$, $p = .007$, BCa 95% CI [.02, .10]. As hypothesized, overall, exposure to a higher proportion of misinformation was associated with a lower proportion of correct memory responses to misinformation items, but this effect was dampened for those participants who were made particularly skeptical of the information source because of high levels of misinformation in the narrative (see Figure 1). The dampening effect can also be seen in Table 2 in the contrast in estimated marginal means and standard errors when credibility estimates are controlled versus not. These results suggest that if the 80% misinformation group had higher credibility estimates (similar to those in the 50% misinformation group), they would have performed with an even lower proportion of correct responses to misleading questions on average, and if the 20% misinformation group had lower credibility estimates they would have performed with a higher proportion of correct responses to misleading questions.

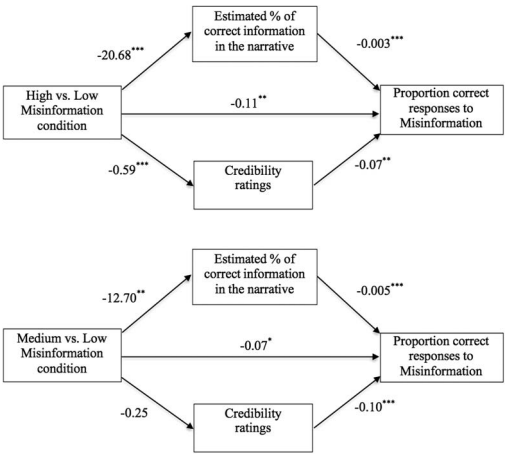

Figure 1. Mediation model reflecting the direct effect of the misinformation condition on the proportion of correct responses to misinformation questions and the mediating roles of the variables reflecting participant perceptions of the narrative. The reported direct effects are independent of the mediator variables. Unstandardized beta values are reported. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 2
Estimated Marginal Means (EMMs) and Standard Errors of Proportions of Correct Responses to Misinformation Questions by Condition Controlling for, and Not Controlling for, Ratings of Credibility and Estimates of the Percentage of the Narrative That Was Incorrect

	Control variables included		Control variables excluded	
Condition	EMM	*SE*	EMM	*SE*
80%	.59	.02	.63	.03
50%	.67	.02	.66	.02
20%	.77	.02	.74	.02

Note. An analysis of covariance with misinformation condition predicting proportion correct, including the control variables, was significant, $F(1, 121) = 15.59$, $p < .001$, $\eta_p^2 = .21$. All differences between misinformation conditions were significant, $p = .014$ (80% vs. 50% condition), $p = .001$ (50% vs. 20% condition), $p < .001$ (80% vs. 20% condition). Analysis of variance results, without the control variables, are presented in the text and in Table 1. Only the 80% and 20% misinformation conditions differed significantly, $p = .006$.

Source-Monitoring and Memory Accuracy

Next, we examined whether skepticism about the credibility of the misinformation narrative was associated with source monitoring. We coded source responses for each memory question as accurate or inaccurate. For example, if a participant selected a response in line with previously presented misinformation and then selected the narrative as the source of their response, the response would be coded as accurate source monitoring. Any other response (i.e., the video, both the video and the narrative, or neither) would be coded as inaccurate source monitoring. Source monitoring accuracy on individual items was averaged to create a source monitoring accuracy variable. As predicted, thinking that the narrative was more credible and was composed of more

correct information was associated with poorer source-monitoring accuracy, $r(125) = -.41$, $p < .001$ and $r(124) = -.41$, $p < .001$.

Metacognition and Memory Performance

Participants were relatively accurate at making global estimates of their performance on the memory questions, except when they had been misled. First, we compared participant estimates of the number of questions they answered correctly to the actual number of questions for which they selected the correct, incorrect, and misled response. Their estimates of the number of questions they answered correctly increased with their actual number of correct responses, $r(124) = .21$, $p = .019$, decreased with the number of incorrect responses, $r(124) = -.22$, $p = .012$, and had no relation with the number of misled responses, $r(124) = .02$, $p = .860$. Next, we calculated a difference score by subtracting the number of questions participants thought they answered correctly from the number of questions they actually answered correctly. A one-sample t test revealed that participants tended to overestimate the number of questions they answered correctly by about two on average, ($M = -2.06$, $SD = 7.42$), $t(125) = -3.12$, $p = .002$. The accuracy of participants' performance estimates tended to increase with actual performance accuracy such that the absolute difference between their actual performance and estimated performance decreased as the number of correct responses increased, $r(124) = -.20$, $p = .023$. Next, we compared participants' estimates of their memory strength to the proportion of misleading items to which they responded correctly. Mean memory strength ratings were approximately at the center of the 7-point scale ($M = 3.58$, $SD = .60$) and increased in line with their actual performance, $r(125) = .24$, $p = .007$. Finally, we assessed the accuracy of participants' performance estimates in terms of confidence in specific responses. An averaged confidence variable was created for participants who responded to all 40 confidence questions (102 participants in total). Higher average confidence estimates ($M = 3.85$, $SD = .55$) were associated with a larger proportion of correct memory responses to misleading items, $r(100) = .21$, $p = .038$. Conversely, confidence decreased as the proportion of incorrect responses to misleading items increased, $r(100) = -.45$, $p < .001$. There was no association between confidence and the proportion of misled responses to misleading items, $r(100) = .03$, $p = .807$.

Discussion

Previous research has demonstrated that people are able to disregard misinformation and more carefully monitor their memories for accuracy when they have reason to believe that the source of the misinformation lacks credibility (Hirst & Echterhoff, 2008; Koppel, Wohl, Meksin, & Hirst, 2014). However, in past research participants are generally provided with cues *external* to the narrative to help them assess source credibility, either before or after they have been provided with misinformation. These cues may not be true reflections of source accuracy in real world contexts (Echterhoff et al., 2005; Smith & Ellsworth, 1987). However, when people can detect discrepancies between their memory for an event and the misinformation presented, they are able to reject misinformation (Tousignant et al., 1986). In the present study, we manipulated the amount of postevent misinformation participants received to examine (a) the association between amount of misinformation and susceptibility to misinformation, (b) whether increasing the proportion of misinformation to neutral information would impact perceived source credibility, (c) whether perceptions of source credibility mediated the misinformation-memory performance association, (d) the relations between perceived source credibility and source monitoring accuracy, and (e) whether participants were able to accurately assess their performance. Our results shed light on human memory processes, including contextual and individual difference predictors of resistance to misinformation, and have important implications for applied contexts such as in clinical settings.

First, we found, unsurprisingly, that exposure to more misinformation led participants to answer a lower proportion of misinformation questions correctly and a higher proportion of misinformation questions in line with the presented misinformation. The high rates of misinformation may have burdened participants with higher cognitive load as they struggled to monitor the high rates of misinformation. This cognitive load may have made them more vulner-

able to the misinformation (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Magliano & Radvansky, 2001).

Critically, in line with our primary hypothesis, an opposing pattern also emerged such that exposure to higher rates of misinformation led participants to be more skeptical of the accuracy of the narrative, and this increased skepticism led to an increase in resistance to misinformation. On average participants who received a narrative with 80% misinformation answered smaller proportions of questions about those misinformation items correctly (i.e., they were less accurate) than those who received 20%. However, within each condition participants performed best if they reported being skeptical of the accuracy of the narrative. Participants who were presented with higher levels of misinformation were most likely to be made skeptical, and those who were made skeptical resisted a greater proportion of presented misinformation.

Skepticism, both in terms of perceived credibility and narrative accuracy, was also associated with increased source monitoring accuracy. This association may have emerged in line with a source monitoring perspective (Koppel et al., 2014), and in line with previous research using explicit warnings about source credibility (Echterhoff et al., 2005; Frost et al., 2002; Skagerberg & Wright, 2009), such that those participants who were distrustful of the misinformation source more vigilantly monitored their memory to resist misinformation. The opposite pattern of results was also possible, such that participants who more vigilantly monitored the narrative for misinformation better detected discrepancies between the video and narrative and thus reported that the narrative was of lower credibility and accuracy. The directionality of this effect needs further exploration in future research.

Participants also demonstrated successful metacognition in terms of estimating their correct and incorrect responses, but not their misled responses. While participants tended to slightly overestimate their accuracy, those who gave more correct and fewer incorrect responses correctly rated themselves as being more accurate and as having stronger memories, and they were more confident in their responses. Participants who answered more memory questions correctly were also better calibrated in their performance estimates; however, the causal direction of this relation is unclear. It may be that all participants rated their performance similarly in terms of accuracy regardless of actual accuracy. If this were the case, estimates would appear more calibrated for highly accurate participants because on average, people estimated that they were highly accurate. Or it may be that participants who were more resistant to misinformation were also better able to estimate their performance. The latter interpretation is in line with research demonstrating that students who perform well on examinations tend to be more attuned to their own performance (see Falchikov & Boud, 1989, for a review). Examination of the present confidence results might shed some light on this. Specifically, subjective confidence ratings were positively related to correctly responding and negatively related to incorrectly responding. Thus, it appears that participants were at least somewhat attuned to how well they remembered the information. In line with previous research, participants appeared unaware when they had incorporated misleading information both in terms of their performance estimates and subjective confidence ratings (Greene, Flynn, & Loftus, 1982; Loftus, Donders, Hoffman, & Schooler, 1989; Luna & Migueles, 2009).

In summary, there appear to be two opposing pathways linking misinformation to memory accuracy: a direct path between provision of misinformation and susceptibility to misinformation and an indirect path via perceptions of source credibility. On the one hand, receiving high proportions of misinformation negatively impacted participants' performance by increasing susceptibility to misinformation. On the other hand, when participants received more misinformation, they were, on average, more skeptical of the source of that misinformation, and appropriately rated the narrative as containing fewer correct details and as being of lower credibility. Critically, this skepticism about the misinformation narrative was associated with *increased* source monitoring accuracy and resulted in a significantly diminished misinformation effect. Thus, this research demonstrates that while provision of misinformation is undoubtedly detrimental to memory performance, individuals who are appropriately skeptical of the accuracy of a misinformation source are partially protected from memory errors.

Our work also has implications for clinical contexts. For example, recovered-memory ther-

apeutic techniques (RMTs) are those in which therapists attempt to extract clients' purportedly "repressed" memories (Poole, Lindsay, Memon, & Bull, 1995). Researchers have revealed that RMTs can involve suggesting event elements not previously mentioned by patients. Moreover, use of these types of techniques in research settings increases the likelihood that individuals will incorporate false details into their memories, or even unintentionally fabricate entire false events (e.g., Hyman, Husband, & Billings, 1995; Loftus & Pickrell, 1995). Thus, RMT has been labeled a potentially harmful psychological technique, both in terms of the potential for memory distortion, and for the mental health of clients (Lilienfeld, 2007). However, according to a recent review article, these techniques still "flourish" (Lynn, Evans, Laurence, & Lilienfeld, 2015). In contrast, empirically supported therapies focus on allowing patients to report and interpret events using their own words, with therapists offering assistance in the deeper exploration and modification of thoughts and feelings related to these experiences, rather than the details of the remembered experiences themselves (Lane, Ryan, Nadel, & Greenberg, 2015). The present research highlights the importance of adhering to these latter practices and avoiding making assumptions, or suggesting additions or revisions to factual content in clinical contexts, particularly given that clinicians are likely perceived as highly credible information sources.

References

Belli, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General, 118,* 72–85. http://dx.doi.org/10.1037/0096-3445.118.1.72

Belli, R. F., Lindsay, D. S., Gales, M. S., & McCarthy, T. T. (1994). Memory impairment and source misattribution in postevent misinformation experiments with short retention intervals. *Memory & Cognition, 22,* 40–54. http://dx.doi.org/10.3758/BF03202760

Chambers, K. L., & Zaragoza, M. S. (2001). Intended and unintended effects of explicit warnings on eyewitness suggestibility: Evidence from source identification tests. *Memory & Cognition, 29,* 1120–1129. http://dx.doi.org/10.3758/BF03206381

Clifasefi, S. L., Garry, M., & Loftus, E. (2007). Setting the record (or video camera) straight on memory: The video camera model of memory and other memory myths. *Tall tales about the mind & brain: Separating fact from fiction* (pp. 60–75). New York, NY: Oxford University Press.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125,* 159–180. http://dx.doi.org/10.1037/0096-3445.125.2.159

Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning and Verbal Behavior, 19,* 695–704. http://dx.doi.org/10.1016/S0022-5371(80)90379-5

Echterhoff, G., Hirst, W., & Hussy, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory & Cognition, 33,* 770–782. http://dx.doi.org/10.3758/BF03193073

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59,* 395–430. http://dx.doi.org/10.3102/00346543059004395

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 552–564. http://dx.doi.org/10.1037/0096-1523.3.4.552

Frost, P., Ingraham, M., & Wilson, B. (2002). Why misinformation is more likely to be recognised over time: A source monitoring account. *Memory, 10,* 179–185. http://dx.doi.org/10.1080/09658210143000317

Greene, E., Flynn, M. S., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior, 21,* 207–219. http://dx.doi.org/10.1016/S0022-5371(82)90571-0

Hirst, W., & Echterhoff, G. (2008). Creating shared memories in conversation: Toward a psychology of collective memory. *Social Research, 75,* 183–216.

Hyman, I. E., Husband, T. H., & Billings, F. J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology, 9,* 181–197. http://dx.doi.org/10.1002/acp.2350090302

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28. http://dx.doi.org/10.1037/0033-2909.114.1.3

Koppel, J., Wohl, D., Meksin, R., & Hirst, W. (2014). The effect of listening to others remember on subsequent memory: The roles of expertise and trust in socially shared retrieval- induced forget-

ting and social contagion. *Social Cognition, 32,* 148–180. http://dx.doi.org/10.1521/soco.2014.32.2.148

Lane, R. D., Ryan, L., Nadel, L., & Greenberg, L. (2015). Memory reconsolidation, emotional arousal, and the process of change in psychotherapy: New insights from brain science. *Behavioral and Brain Sciences, 38,* e1. http://dx.doi.org/10.1017/S0140525X14000041

Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2,* 53–70. http://dx.doi.org/10.1111/j.1745-6916.2007.00029.x

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7,* 560–572. http://dx.doi.org/10.1016/0010-0285(75)90023-7

Loftus, E. F. (1979). Reactions to blatantly contradictory information. *Memory & Cognition, 7,* 368–374. http://dx.doi.org/10.3758/BF03196941

Loftus, E. F., Donders, K., Hoffman, H. G., & Schooler, J. W. (1989). Creating new memories that are quickly accessed and confidently held. *Memory & Cognition, 17,* 607–616. http://dx.doi.org/10.3758/BF03197083

Loftus, E. F., & Hoffman, H. G. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General, 118,* 100–104. http://dx.doi.org/10.1037/0096-3445.118.1.100

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19–31. http://dx.doi.org/10.1037/0278-7393.4.1.19

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25,* 720–725. http://dx.doi.org/10.3928/0048-5713-19951201-07

Luna, K., & Migueles, M. (2009). Acceptance of central and peripheral misinformation and confidence. *The Spanish Journal of Psychology, 12,* 405–413. http://dx.doi.org/10.1017/S1138741600001797

Lynn, S. J., Evans, J., Laurence, J. R., & Lilienfeld, S. O. (2015). What do people believe about memory? Implications for the science and pseudoscience of clinical practice. *Canadian Journal of Psychiatry, 60,* 541–547. http://dx.doi.org/10.1177/070674371506001204

Magliano, J. P., & Radvansky, G. A. (2001). Goal coordination in narrative comprehension. *Psychonomic Bulletin & Review, 8,* 372–376. http://dx.doi.org/10.3758/BF03196175

Otgaar, H., Candel, I., Scoboria, A., & Merckelbach, H. (2010). Script knowledge enhances the development of children's false memories. *Acta Psychologica, 133,* 57–63. http://dx.doi.org/10.1016/j.actpsy.2009.09.002

Poole, D. A., Lindsay, D. S., Memon, A., & Bull, R. (1995). Psychotherapy and the recovery of memories of childhood sexual abuse: U.S. and British practitioners' opinions, practices, and experiences. *Journal of Consulting and Clinical Psychology, 63,* 426–437. http://dx.doi.org/10.1037/0022-006X.63.3.426

Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. Boston, MA: Houghton, Mifflin and Company.

Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin & Review, 1,* 357–375. http://dx.doi.org/10.3758/BF03213977

Skagerberg, E. M., & Wright, D. B. (2009). Susceptibility to postidentification feedback is affected by source credibility. *Applied Cognitive Psychology, 23,* 506–523. http://dx.doi.org/10.1002/acp.1470

Smith, V. L., & Ellsworth, P. C. (1987). The social psychology of eyewitness accuracy: Misleading questions and communicator expertise. *Journal of Applied Psychology, 72,* 294–300. http://dx.doi.org/10.1037/0021-9010.72.2.294

Takarangi, M. K. T., Parker, S., & Garry, M. (2006). Modernising the misinformation effect: The development of a new stimulus set. *Applied Cognitive Psychology, 20,* 583–590. http://dx.doi.org/10.1002/acp.1209

Tomes, J. L., & Katz, A. N. (2000). Confidence–accuracy relations for real and suggested events. *Memory, 8,* 273–283. http://dx.doi.org/10.1080/09658210050117708

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition, 14,* 329–338. http://dx.doi.org/10.3758/BF03202511

Weingardt, K. R., Leonesio, R. J., & Loftus, E. F. (1994). Viewing eyewitness research from a metacognitive perspective. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 157–184). Cambridge, MA: The MIT Press.

Weingardt, K. R., Loftus, E. F., & Lindsay, D. S. (1995). Misinformation revisited: New evidence on the suggestibility of memory. *Memory & Cognition, 23,* 72–82. http://dx.doi.org/10.3758/BF03210558

Zaragoza, M. S., & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 934–945. http://dx.doi.org/10.1037/0278-7393.20.4.934

Received February 8, 2017
Revision received August 30, 2017
Accepted August 31, 2017 ■